Integrating Perceptual and Symbolic Information in VR

Jay Bolter
School of Literature, Communication & Culture

Larry F. Hodges
College of Computing

Thomas Meyer
College of Computing

and

Alison Nichols
School of Literature, Communication & Culture

Graphics, Visualization & Usability Center
Georgia Institute of Technology (Georgia Tech)
Atlanta, GA 30332
http://www.cc.gatech.edu:80/gvu/virtual/Venue/

To Appear in *IEEE Computer Graphics & Applications* VR Blackboard (July 1995)

Sponsored by Hewlett Packard

INTEGRATING PERCEPTUAL AND SYMBOLIC INFORMATION IN VR

Jay Bolter, Larry F. Hodges, Thomas Meyer and Alison Nichols
Graphics, Visualization & Usability Center
Georgia Institute of Technology (*Georgia Tech*)

What is virtual reality (VR) good for and what are it limits? Many enthusiasts think of virtual reality as the ultimate computer application. They suggest that our society will soon be using VR for a wide range of purposes from recreation to communication to scientific and medical research. The enthusiasts sometimes talk as if immersive virtual reality will be the environment of choice for general computing. In fact, only a few VR applications have progressed beyond the laboratory to real business or popular uses. Stephen Ellis claimed last year that in the past thirty years the only major commercial product to emerge from virtual environment or teleoperation technology is flight simulation. (*IEEE Computer Graphics & Applications,* January 1994, p 21)

Inside or outside the lab, most current VR aims at reproducing perceptual experience. The goal is to give the user the sensation that he or she is walking along a hallway in a building, flying over a city, or riding in a vehicle. However, if virtual reality is ever going to "break out" into general purpose computing, then it has to provide more than a perceptual experience. It has to reproduce the capabilities of current computers to manipulate symbols as well. Symbol manipulation is an essential part of what computers do in applications ranging from numerical analysis to database management to word processing and communication.

In the commercially successful virtual application, flight simulation, the numerical and other symbolic data are at least as important as the perceptual experience. Furthermore, the symbolic data and perceptual experience are relatively easy to separate. The computer creates a graphic scene outside the cockpit window for the pilot, but the instruments are physical devices placed exactly as they would be in a real cockpit. In flight simulation the relationship of text, graphics, and perceptual simulation are demarcated by the nature of the task.

The problem of integrating symbolic and perceptual information becomes more difficult in a fully immersive virtual reality, in which the whole visual apparatus must be drawn and rendered by the computer in real time. One obvious reason that text and numerical information is not incorporated in many virtual environments is the technical limitations of the hardware. The resolution of many head-mounted displays is barely adequate to display readable text. As a result, most VR applications limit themselves to the display of a few "huge" letters or numbers for orientation or navigation. There is also perhaps a lingering prejudice that VR should be able to convey information through its interactive and three-dimensional visual display alone; it should not need to rely on "old-fashioned" representation, the traditional coded forms of numbers and words.

We believe, however, that virtual environments should convey the widest possible range of perceptual and symbolic information. We are therefore interested in testing the informational limits of immersive environments. The first question to ask is whether certain kinds of symbolic information can be conveyed in virtual reality at all. And even if symbolic information can be conveyed,, we still have to consider when virtual reality is worth the trouble. In which application areas does VR give the user increased capabilities that justify the expense and inconvenience of immersive hardware and software? When is it better to apply more mundane 2D (or 2 1/2 D) solutions? These are user interface questions.

VR as a User Interface

There is a widespread conviction that VR should require as little interface as possible. VR is supposed to eliminate the need for anything that mediates between the user and the information. The argument goes as follows. In a virtual environment, the user should be able to use the same tools for interaction that he or she uses in the real world -- eyes and hands (or even the whole body, if a body suit is worn). The user should be able to employ these tools to navigate in and manipulate the environment with motions appropriate to eye and hand movements in the real world. The user does not need to master an elaborate and artificial set of tools, as she does, for example, with the desktop metaphor.

This concept of unmediated interaction may or may not work in virtual environments that are wholly perceptual, like architectural walkthroughs. If we just want to reproduce some portion of an imagined physical world for the user to inhabit, then it is possible to minimize the interface. Even here, however, we may need to introduce such "unnatural" devices as flying by means of a hand gesture. In any case, unmediated interaction is clearly not sufficient when we introduce symbolic information, such as texts, numbers, graphs, or even graphics used symbolically (as in scientific visualization). This information requires an interface. We wish to explore virtual reality as a conscious and elaborate interface, as a highly artificial metaphor for combining symbolic and perceptual information to the user.

Virtual Athletic Venues

To explore the possibilities of such an interface, we are creating two virtual athletic venues. One site is a natatorium, an arena for water sports; the second is a stadium for track and field events. In the natatorium we will stage a diving event and in the stadium pole-vaulting and other track and field events. Our VR users will be the spectators at these sites. They will have the opportunity to walk or fly around the venues, watch the athletes perform, and gather various kinds of information about the sports activities they see.

We are using athletic venues as our model environment, but we intend to develop interface components that will apply to other kinds of virtual models. We are not creating an information kiosk. A kiosk is designed for a casual user who comes up to the computer, interacts for a few minutes, and then leaves. The user of a kiosk wants specific information for a relatively well-defined and limited purpose. Our goal is to present information in greater depth. Our imagined audience is not a casual visitor, but rather a user with considerable interest in the event or the history of athletics. We assume such a user precisely because this assumption give us greater freedom to explore sophisticated interface devices. Our interface may require some training, but the training will be justified by the fact that the user will make extended use of the system.

Some virtual environments, particularly architectural walkthroughs, are relatively static. Sports venues offer us the chance to focus the user's attention on elements that change over time (the athletes performing), and this focus allows us to build an interface that controls both how the athlete is viewed and how the athlete moves. Another reason for our choice is that virtual athletic venues allow us to experiment with the presentation of at least three distinct kinds of information.

Visual Presence and Point of View

One kind of information is simply the experience of being in the venue and viewing the event from a variety of perspectives. VR gives the user the sense of "being there": it creates this sense of presence primarily by allowing the user to control her point of view. Any one frame in a VR scene is not very convincing. The image has a cartoon-like character. Texture mapping can make the image much more lifelike. But a VR image cannot compete in photorealism with a high-quality computer graphic not rendered in real time, and of course in most cases no computer-generated graphic can compete with a photograph. VR makes up for the lack of quality of each single frame by letting the user's head movements determine the perspective. The fact that the scene responds to the user's movements creates a sense of three-dimensional space. The standard interface elements here include tracking head movements and allowing some method of moving through the space. Our environment of course includes both these elements. In addition, we want to enhance the user's control over point of view.

The primary object of interest for atheletic venues is the performing athlete. Currently, we have used modeling software to create relatively crude animations for an athlete in the pole vault. Our next step will be to use motion capture or simulation data to create smooth responsive models of athletes in diving, vaulting and possibly some running events. These models will then perform their dives and vaults as the spectator watches. Obviously the spectator is free to fly around the environment and view the action from any perspective. We also allow the user to link her perspective to the action. For example, the user can take up a viewing position and then synchronize her movement with the vertical or horizontal motion of the athlete. Thus, the user can occupy a position at the same height as the diving board and then lock her vertical motion to that of the diver. As the diver dives, the user will sink down into the pool alongside the diver and can examine the diver's movements from this unique perspective.

We believe that synchronizing one's perspective with a moving object may be useful in a variety of VR applications, including scientific visualization when the scientific user wants to examine data that change over time.

Numerical/Statistical information

Athletics is about names and numbers: who ran the fastest? who scored the most points? how many contests were won by a particular athlete over a lifetime? And so on. We want to make that kind of information visible and manipulatable. One technique is to overlay numbers onto the virtual scene. For example, as the diver dives, the user will be able to see the rate of descent as a number that floats in her field of view. The effect will be rather like that used in such science fiction films as the Terminator. Although our environment is not a kiosk, we may also include information kiosks within the venue to display numerical information. The user interface issues here are both how the numbers can best be displayed and how the user asks for and adjusts the display within the environment.

Another way to provide some numerical information is to impose a grid behind the athlete to allow the user to measure his motion. The grid may supplement or replace a display of digits in the user's field of view. Gridding has the advantage that it conveys numerical information without the use of digits, which may be hard to read with a head-mounted display.

The display of such information is obviously related to the use of VR for scientific visualization. The scientific user may want to see the numbers associated with the some aspect of the data. The capacity for effective display of numbers and the use of grids may make VR a more successful environment for scientific visualization.

Narrative Information

Athletic competition is also a story that can be told in words or in sound and images. The narrative can, for example, give the history of a particular event within modern athletics, stories associated with a particular athlete, or techniques for diving or vaulting. Our virtual venues will therefore allow for the display of verbal text, audio, and even digitized video.

Verbal information can be presented as speech or as visible text. We hope ultimately to provide visible text, but current limitations on resolution make it difficult to display large amounts of readable text in virtual reality. For that reason, we are experimenting first with spoken text. The user will be able to click on various annotation markers around the venue to receive digitally recorded sound. The interface required is part of a more elaborate system for making and retrieving annotations. In a full system, the user should be able to annotate objects in the environment as well as particular spatial or temporal views. These annotations can function as hyperlinks: they allow the user to "branch" to material that explains or elaborates on what she is currently seeing. We believe that annotation can be useful in a variety of virtual environments, particularly for scientific visualization. A scientific user could make notes without having to leave the environment. The notes could relate to specific data elements or to spatial or temporal slices of the data. The notes could be replayed by collaborators or by the same user on a later visit to the data.

Digitized video or film constitutes another and very different medium for narration. Unlike VR, film or video unrolls at a predetermined rate and with a point of view not controlled by the viewer. Therefore film and video can communicate an author's point of view and information in a way that VR cannot. As a first test, we will present in the natatorium footage of the 1936 athletic diving from the Riefenstahl film Olympia. By placing such a conventional film or video into the virtual environment, we are effectively making VR into a site for multimedia.

Another challenge is to devise an interface for showing and controlling the movie. The movie itself will be available on a screen in the venue. When the user activates the screen, the movie will blow up to occupy her whole field of view, except possibly for a small control area. When the movie has finished playing, the movie image shrinks and the user is returned to the venue.

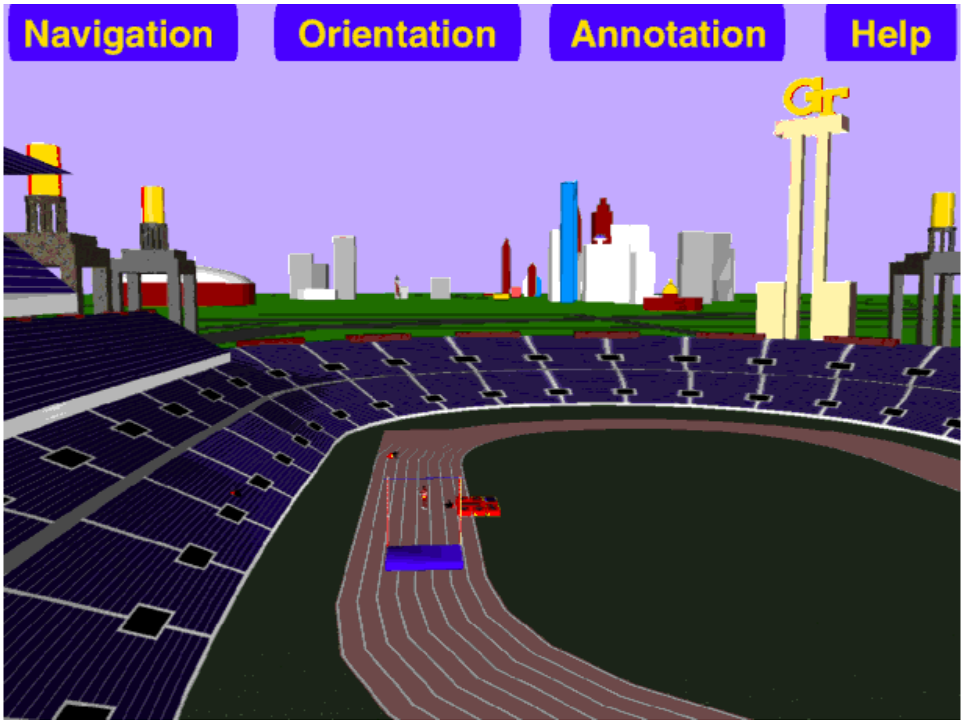
In our Virtual Venues, each film clip will function as a multimedia annotation, providing background and context for the event. Again, we believe that an effective interface for the presentation of video annotations is a tool that has broad application in virtual environments.

Heterogeneous Virtual Reality

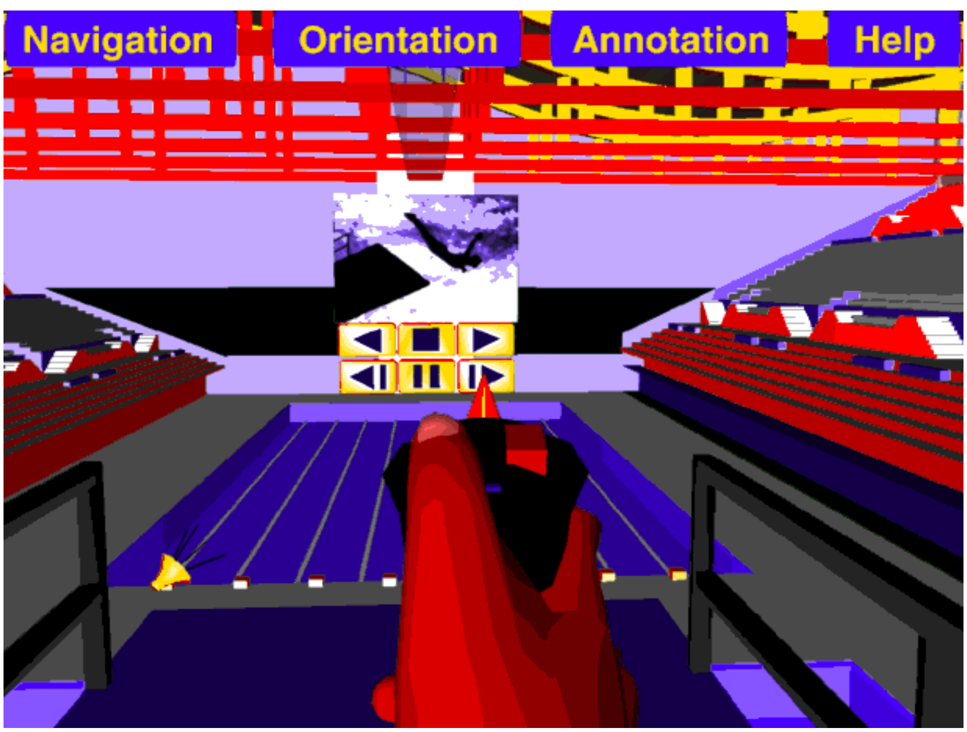
Virtual Athletic Venues is an experiment in what we might call "heterogeneous virtual reality." Two or three decades ago, most computing applications were homogeneous. Each served one purpose, or a few related purposes, and each had one preferred mode of input and output. Today most computing applications are bundled into environments consisting of many different facilities and often gather and present information in a variety of different media. Virtual reality has not reached this stage of maturity. However, if VR is ever to be more than a highly specialized research tool, it must provide the heterogeneity that users have come to expect from other environments. VR must be able to handle the various media that now make up the digital information revolution.

Acknowledgments

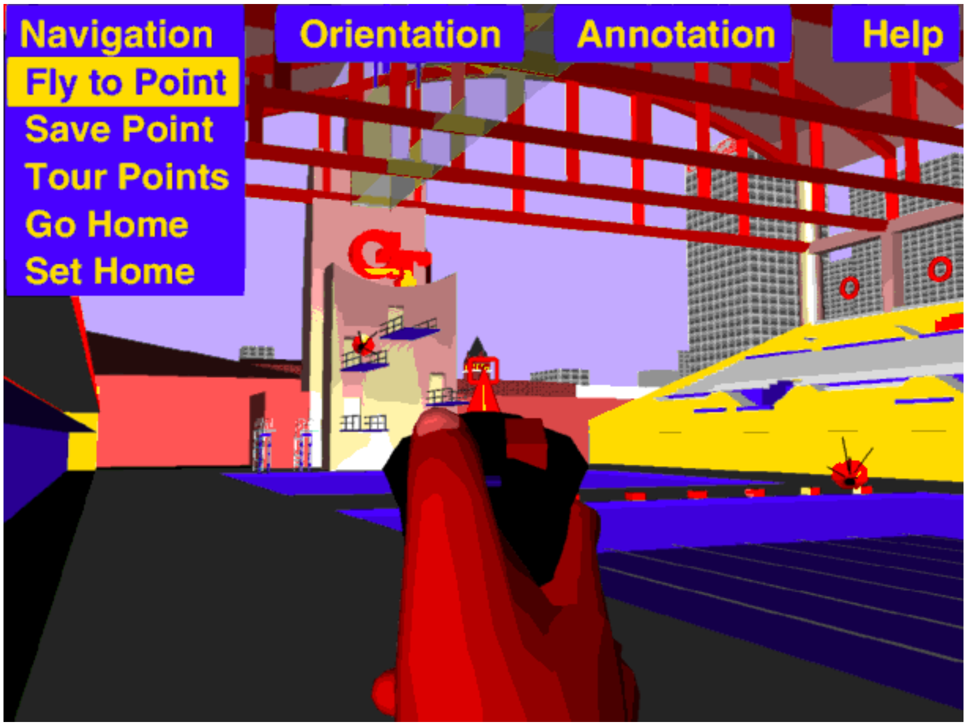
Hewlett Packard has provided generous support for this project in the form of both student research assistantships and HP graphics workstations.



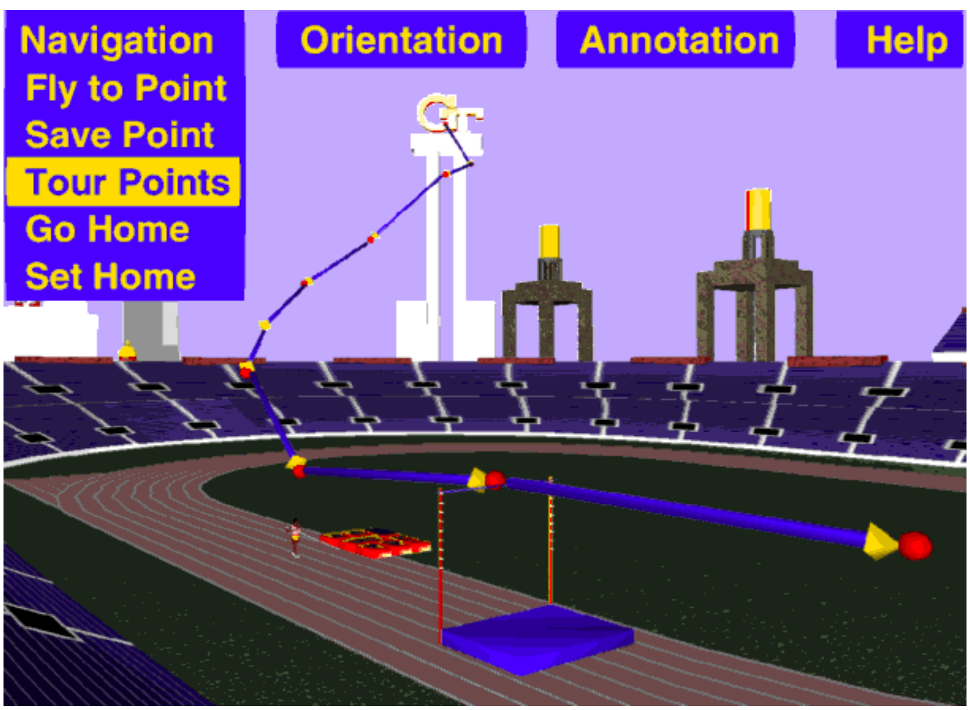
Atlanta's Track and Field Stadium: Pull-down menu items for Navigation, Orientation, Annotation, and Help functions can be operated using a 3-D pointer in the HMD/Tracker environment or with the mouse in the screen-based environment.



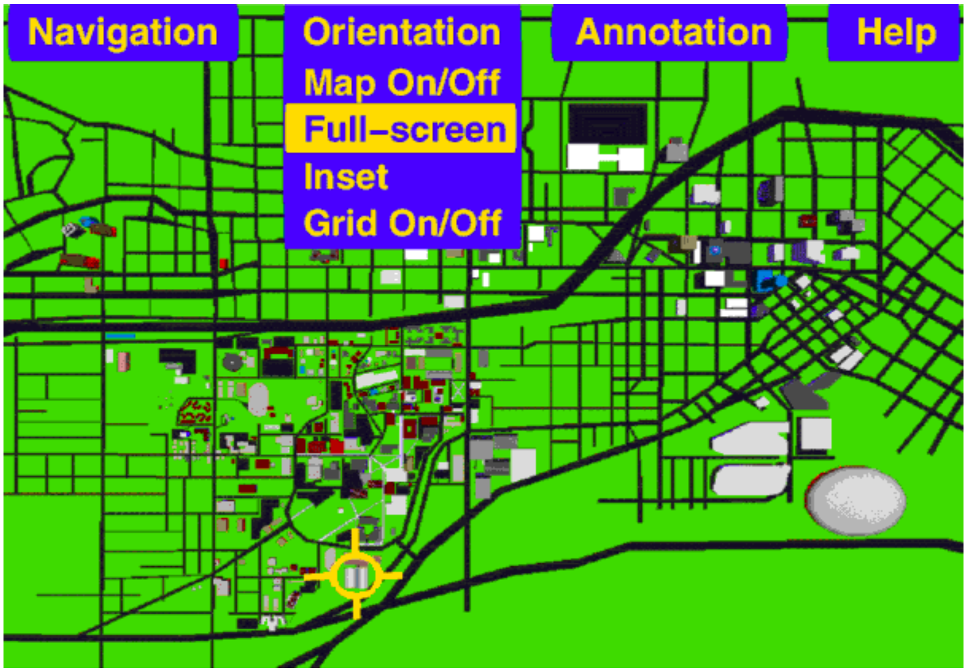
Georgia Tech's Natatorium: Video annotations projected onto a screen can be played back by selecting menu buttons with a 3-D pointer



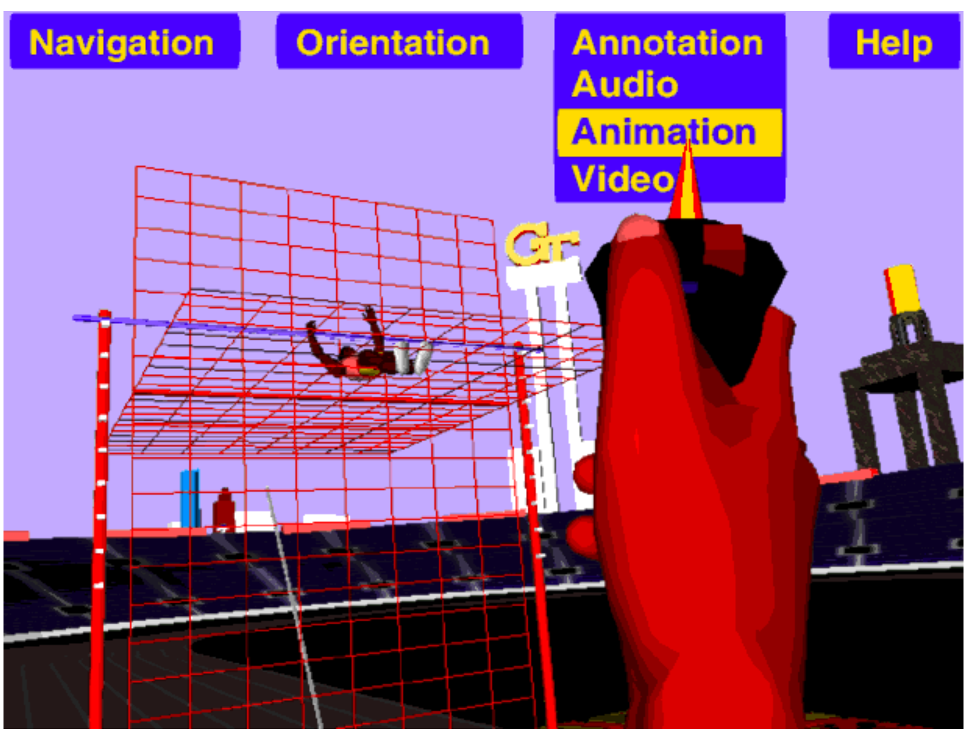
Natatorium Diving Platform: Navigating to distant objects with the "Fly-to-point" option–just point to an object with the 3-D pointer and click a button.



Track and Field Stadium: A path of points can be recorded in the environment and played back for a tour. The tour path is shown in the environment as a series of 3-D links between saved points.



Bird's Eye View of the Model of Atlanta and the Georgia Tech Campus: As an orientation aide, a full-screen or inset map can shown to indicate the user's current location. As the user navigates the environment, a circle around the user's location is reprojected onto the static map.



Animated Pole Vaulter: Audio, video, and animation annotations can be placed into the environment for later playback. Annotations are indicated by 3D annotation markers which appear in the environment, and each media type includes its own buttons for playback of the content.